More Expression, Less Error Alternative Ballots Outperform Status Quo

Jason Stuart Maloy, University of Louisiana at Lafayette

Overview

This research brief reports results from online experiments conducted at the height of the 2020 U.S. presidential primary season. Over 6,000 experimental subjects in four American states which held their primary elections on Super Tuesday—Colorado, Tennessee, Texas, and Virginia—voted in simulated elections involving real presidential candidates and three different ballot types: a traditional single-mark ballot, a ranking ballot, and a grading ballot.

Research Questions

1. Which ballot types are more likely in general to produce void votes (ones that cannot be counted toward the final result)?
2. Which types of voters are more likely in general to cast void votes?
3. How do different ballot types affect the likelihood that some types of voters will cast more void votes than others?

Key Findings

1. Alternative ballot types (ranking and grading) produced fewer void votes (more valid votes) than the status-quo ballot type (single-mark).
2. BIPOC voters and male voters were more likely to cast void votes, but these discrepancies were moderated by past voting experience and experience in higher education.
3. Alternative ballot types (ranking and grading) were associated with smaller discrepancies in error-proneness according to race and gender, while the status-quo ballot (single-mark) was associated with larger discrepancies.

Background and Research Design

Voting access is a crucial factor in the quality of electoral democracy in any city, state, province, or nation. Accordingly, people are rightly concerned about their ability to register as eligible voters and to exercise their right to vote for governmental officers in the jurisdictions to whose laws they are subject. At the same time, there is an important second-order issue for voting access: Once you're allowed into the voting booth, how easy or hard is it for you to cast a valid vote that will count toward the outcome?

This second issue, the accessibility of the ballot itself, has been amply explored in terms of voting technology, ballot design, and administrative procedure since the notorious presidential recount of 2000 in the state of Florida. Much money has been spent, and some improvements have been made, in most American states since that time.

But what about the voting method itself—the input rules that voters are given in the instructions on the ballot—and its implications for ballot accessibility? An alternative voting method known as ranked-choice voting (RCV), where voters rank the candidates in order of preference, has been implemented in citywide elections in 18 cities around the country since 2006 and in statewide elections in Maine since 2018. Other reform options such as "range" or "evaluative" voting—essentially a grade point average (GPA) type of system but applied to political candidates instead of students in school—have been proposed but not implemented.

The common intuition is that no input rules could be simpler than the traditional, single-mark ballots that are already used in most American elections. Pick a single favorite candidate for each office and make a single mark to indicate your vote for that favorite. By contrast, asking voters to make rankings or to give grades to more than one candidate per contest seems complicated, inviting an explosion of voting error. In other words, it's easy to assume that there's a trade-off between the accessibility of the voting method and how expressive it is: All-or-nothing is crude but accessible, whereas giving different degrees of support to different candidates is more expressive but also more error-prone.

Is the conventional wisdom actually true? To find out, we have to investigate how different voters use different voting methods. Voting experiments are a vital tool for addressing this question, for two reasons. First, for proposals that have never been implemented, such as GPA voting, we have no data from real elections to go by. But experimental studies can generate such data, which are useful as long as we note the features of the experimental setting that differ from those of a real election. (For example, in real public elections with either optical-scan or touch-screen voting machines, the voter may be alerted if some error is detected; in the "error-friendly" design for these experiments, no error alert was used.) Second, even for a voting method like RCV that has an actual track-record (albeit a relatively brief one), we can't see who exactly is having trouble and who isn't without violating the secrecy of the ballot. In a voting experiment, however, we can anonymously collect information on individual voters' characteristics while correlating it with what they do inside the experimental voting booth.

With this goal of analyzing voting error at the individual level in mind, we conducted online voting experiments involving real presidential candidates' names in four American states in February and March of 2020. Over 6,000 respondents voted in a simulated Democratic Party presidential primary and in a simulated blanket (all-party) presidential primary contest.

The ballot type for each voting task was randomly assigned from three options, called "check," "rank," and "grade." The check ballot mimicked the traditional, single-mark ballot, on which the voter is instructed to check a box next to one and only one candidate's name. The rank ballot offered a range of rankings (first, second, third, etc.) next to each candidate's name, instructing the voter to give a unique ranking to as many candidates as desired. The grade ballot offered a range of grades with point-scores (A for four points, B for three, C for two, etc.), instructing the voter to select any one grade for any candidate (with F for zero points as the default).

Results and Implications

The main expectation behind the experiments was that the simplest and most familiar ballot type, the check ballot, should produce the lowest proportion of "void" votes, or votes that cannot be counted toward the final result. Yet results show that the traditional ballot produced more void votes than the reform alternatives (Figure 1). At the same time, the rates of mismarked votes, with at least one error on a completed ballot, were higher for the reform alternatives than for the traditional ballot. Other results suggest that group-based inequalities in voting error—measured by the discrepancies in void rates across age, gender, and race cohorts—are smaller with reform ballots than with the status quo (Figure 2).

Table 1 | Three aspects of voting error by ballot type

Ballot Type	Blank	Invalid	Mismarked	Void Rate
Check	1.1%	6.9%	6.9%	8.0%
Rank	0.8%	5.7%	15.4%	6.5%
Grade	0.4%	5.4%	8.8%	5.8%

*For the Rank and Grade treatments, a "mismarked" ballot was considered valid if at least one clear candidate preference was indicated. For the Check treatment, the input rules and counting rules require every mismarked ballot to be considered void.

NEW AMERICA

Download image

Figure 1 (above) allows comparison of different kinds of voting error across the three ballot treatments. The "blank" rate measures those ballots that contained no markings for any candidate. The "invalid" rate includes those ballots that were marked in such a way that they could not be added to the count. The "mismarked" rate measures those ballots that contained at least one violation of the instructions. Whether a mismarked ballot could still be counted depended on the ballot type. For the rank and grade treatments, a mismarked ballot was still considered valid if at least one clear candidate preference was indicated. For the check treatment, as for single-mark ballots in the real world, the input rules and counting rules require every mismarked ballot to be considered void.1

These results suggest that the opportunity to mark each and every candidate on ranking and grading ballots generates more violations while nonetheless allowing more voters to express a clear preference for at least one candidate. Figure 2 (below) shows the effect of different ballot treatments on different racial cohorts. For BIPOC voters in general, the opportunity to use ranking or grading ballots to express their electoral judgments, instead of traditional single-mark voting, tended to put them at less of a disadvantage with respect to their white peers.

Table 2 | Void rates by racial cohort and ballot type

Ballot Type	White	Black	Latino	Asian
Check	5.6%	14.2%	11.6%	15.8%
Rank	5.1%	11.9%	7.1%	7.0%
Grade	4.8%	6.9%	10.8%	8.3%

NEW AMERICA

The results reported here puncture the myth of simplicity surrounding single-mark ballots, the status quo in most American elections. In the 2020 Super Tuesday experiments in four states, the status quo was not effectively exploited as a more natural or obvious way of voting, compared to the alternatives. Whether any given ballot type is easy or hard to use correctly cannot be inferred from its formal input rules alone, and it may be in part a function of patterns of political socialization.

Read more in Jason Maloy Voting Error Across Multiple Ballot Types:Results from Super Tuesday (2020 Experiments in Four American States, **SSRN, October 5, 2020**